The Factor Structure and Measurement Invariance of Positive and Negative Affect

A Study in Iran and the USA

Mohsen Joshanloo and Ali Bakhshi

Department of Psychology, Keimyung University, Dalse-gu, South Korea, Department of Psychology, Qazvin Islamic Azad University, Iran

Abstract. This study investigated the factor structure and measurement invariance of the Mroczek and Kolarz’s scales of positive and negative affect in Iran (N = 2,391) and the USA (N = 2,154), and across gender groups. The two-factor model of affect was supported across the groups. The results of measurement invariance testing confirmed full metric and partial scalar invariance of the scales across cultural groups, and full metric and full scalar invariance across gender groups. The results of latent mean analysis revealed that Iranians scored lower on positive affect and higher on negative affect than Americans. The analyses also showed that American men scored significantly lower than American women on negative affect. The significance and implications of the results are discussed.

Keywords: positive affect, negative affect, factor structure, measurement invariance, cultural differences, gender

Positive and negative affect are two components of subjective well-being, with the other component being life satisfaction (Busseri & Sadava, 2011). Prior research has generally shown that positive and negative affect are separable components rather than two ends of a single continuum (for a review see Schimmack, 2008). There are many factors that play a role in affective experience, including culture and gender. Cultures vary in various aspects of affective experience, including duration, intensity, and expression of emotional experience, as well as the types of emotion favored, and emotion regulation styles (e.g., Fischer & Manstead, 2000; Mesquita & Albert, 2007). Most relevant to the focus of the present study is that cultures tend to vary in the frequency of negative and positive emotions experienced. Research indicates that, in general, individuals in Western individualistic cultures experience positive emotions more frequently, and negative emotions less frequently, than individuals in non-Western collectivistic cultures (e.g., Diener & Suh, 1999; Kuppens, Ceulemans, Timmerman, Diener, & Kim-Prieto, 2006; Lischetzke, Eid, & Diener, 2012).

The differences between these cultures in the frequency of experienced emotions can be partly explained by economic development (Diener & Suh, 1999). In wealthy nations (such as North American countries), satisfaction of physical and social needs of individuals is facilitated, which leads to higher subjective well-being (Tay & Diener, 2011). Religion also plays a role in creating these differences between individualistic and collectivistic nations. Many individualistic nations (including Western European and North American nations) are influenced by Christianity, and many collectivistic nations (such as Muslim and East Asian nations) are influenced by Islam, Buddhism, and Hinduism. Religion is an important part of culture, which influences emotional experience (Watts, 1996). For example, Muslims have been found to experience more sadness and shame than Christians (Kim-Prieto & Diener, 2009).

Gender also affects the frequency of experienced emotions. Numerous studies in Western samples have found that women tend to score higher than men on the frequency of negative emotions, such as sadness and anxiety (for a review see Nolen-Hoeksema & Rusting, 2003). With regard to positive emotions, however, research has yielded inconsistent findings (for a review see Nolen-Hoeksema & Rusting, 2003). However, the general finding is that women and men do not differ significantly on positive affect (e.g.,
Thompson, 2007). Diener, Tay, and Myers (2011), for example, found the relationship between gender and positive affect to be zero in a large sample from 154 nations.

Positive and negative affect have been measured using various scales, including the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988), which is widely used in emotion and well-being research. Mroczek and Kolarz (1998) have developed another brief instrument to measure positive and negative affect. These affect scales have yielded alphas of 0.91 and 0.87, respectively, in a sample of 2,727 American adults (Mroczek & Kolarz, 1998). These scales have been used in several studies with Iranian university students (e.g., Joshanloo, 2011; Joshanloo & Ghaedi, 2009), yielding internal consistency generally higher than .80. Bakhshi, Daemi, and Ajilchi (2009) established convergent validity of these scales through significant correlations with psychological and social well-being in an Iranian sample. The scales have also been found to correlate with other well-being measures in American samples (e.g., Keyes, Shmotkin, & Ryff, 2002; Walen & Lachman, 2000), indicating acceptable convergent validity.

The Present Study

Although Mroczek and Kolarz’s scales have been used in many studies, and there is promising evidence on their convergent validity and reliability, not much is known about their factor structure, measurement invariance, and cross-cultural validity. The present study sought to fill this gap using two large samples from Iran and the USA. These nations have notable differences in terms of language, religion, geography, and socioeconomic indicators. For example, they differ on individualism, with Iran scoring remarkably lower than the USA (Hofstede, Hofstede, & Minkov, 2010). They also differ in terms of the human development index (measuring life expectancy, education, and income), with the USA scoring considerably higher than Iran (UNDP, 2012). Whereas Iranians are dominantly influenced by Islam, American culture is historically dominated by Christianity, and there is a larger religious diversity in the contemporary USA compared to Iran. With these remarkable differences, Iran and the USA provide a good opportunity to test the cross-cultural validity of Mroczek and Kolarz’s affect scales. Toward this goal, the present study investigated the factor structure and measurement invariance of the scales in these two nations. In addition, given that gender has been found to influence emotional experience, measurement invariance of the scales was investigated across gender groups.

Another aim of the present study was to investigate latent mean differences (Byrne, Shavelson, & Muthén, 1989) in positive and negative affect, between national and gender groups. In view of the studies showing that wealthy individualistic countries tend to score higher on positive affect and lower on negative affect than non-Western collectivistic countries (e.g., Diener & Suh, 1999), we predicted that the USA would score higher on positive affect and lower on negative affect. This prediction is also supported by a number of prior studies that have found Iran to score lower than the USA on subjective well-being (e.g., Kuppens et al., 2006; Minkov, 2009). With regard to latent mean differences between gender groups, given the evidence reviewed above, we predicted that women would score higher on negative affect, but there would be no gender difference in the frequency of positive affect.

Method

Measure

Mroczek and Kolarz’s (1998) positive and negative affect scales were used. The scales have been translated into Persian using the method of back-translation (e.g., Joshanloo, 2011; Joshanloo & Ghaedi, 2009). Respondents indicated how much of the time—all, most, some, a little, or none of the time—during the past 30 days, they felt six positive and six negative affective states. Responses were recoded such that higher scores indicated higher frequency of the experienced affect.

Participants

Iranian Sample

The Iranian data was obtained from the Iranian Mental Well-being Project (IMWP) which has been run by the first author in the period of 2003 to the present. This large sample has been collected using convenience sampling, and includes 1,966 students studying at universities in Tehran, 221 adults, 89 participants from a young religious group, and 125 Zoroastrian Iranians (total N in the present study = 2,401). However, 10 participants with missing values on more than 3 out of 12 affect items were excluded, including 8 participants with missing values on all the items, and two participants with 4–6 missing values. Therefore, the final sample size for Iran was 2,391. Females constituted 47.9% of the sample. The mean age was 23.83 (SD = 7.230).

US Sample

The US data was obtained from the second wave of the National Study of Midlife in the United States (MIDUS II; Ryff, et al., 2012). Data collection took place in 2004–2006. The random digit dialing sample (N = 2,257) and oversamples from five metropolitan areas in the USA (N = 489) were used (N = 2,746). However, 592 participants with missing values on more than 3 out
of 12 affect items were excluded, including 567 participants with missing values on all the items and 25 participants with 4 to 11 missing values, leaving a final sample of 2,154. Females constituted 53.1% of the sample. The mean age was 56.73 (SD = 12.623).

**Statistical Analysis**

For testing measurement invariance in multigroup confirmatory factor analysis, each group is individually tested for goodness of fit. Then a series of increasingly restrictive measurement invariance tests are performed. The most important tests are for configural, metric, and scalar invariance (e.g., Chen, 2008; Cheung & Rensvold, 2002), which are conducted here. In the present study, the models were compared using ΔCFI test. An absolute difference in CFI that is less than 0.01 (ΔCFI < −0.01) would indicate measurement invariance (Cheung & Rensvold, 2002). The analyses were conducted using Mplus 6. Full information maximum likelihood (FIML) was used for handling missing data.

**Results**

**Analyses Across Cultures**

**Single-Group CFAs**

A single-factor model of affect was tested separately in each country. This model provided a poor fit to the data both in Iran ($\chi^2 = 3,208.297$, $df = 54$, $p < .001$, $CFI = 0.74$, $RMSEA = 0.156$) and the USA ($\chi^2 = 3,562.310$, $df = 54$, $p < .001$, $CFI = 0.77$, $RMSEA = 0.174$). A two-correlated-factor model resulted in much better fit in Iran ($\chi^2 = 1,383.245$, $df = 53$, $p < .001$, $CFI = 0.89$, $RMSEA = 0.102$) and the USA ($\chi^2 = 1,512.506$, $df = 53$, $p < .001$, $CFI = 0.90$, $RMSEA = 0.11$). This indicates that a two-factor model is more consistent with the data than a single-factor model and, thus, in all the subsequent analyses in the present study, positive and negative affect are considered to be two correlated factors.

The modification indexes in both countries indicated that specifying covariances between the residuals of items 2 and 3 of the negative affect scale, and items 1 and 2 of the positive affect scale would significantly improve the fit of the two-factor model in both countries. Specifying these item residual covariances resulted in excellent fit in both countries. The fit indexes are presented in Table 1. The parameters of the two-factor model with residual covariances are reported in Table 2, separately for each country. This model was used as the baseline model in the subsequent analyses. The latent correlations between the two factors were −0.65 and −0.67 in Iran and the USA, respectively. The manifest correlations were −0.53 and −0.60 in Iran and the USA, respectively (all four correlations were significant at $p < .001$).

**Measurement Invariance**

The baseline model with no equality constraint was simultaneously tested across the two groups. As can be seen in Table 3 (M1), the fit of this model to the data was acceptable, indicating that configural invariance is supported. Equality constraints were then imposed on all factor loadings across the countries. As shown in Table 3 (M2), the ΔCFI ($= −0.008$) indicated full metric invariance. Next, equality constraints were imposed on all item intercepts to test scalar invariance (M3). The ΔCFI ($= −0.024$) indicated noninvariance. Inspection of the modification indexes suggested that freeing the constraints for two items of the positive affect scale (items 2 and 6) would improve the fit of the model substantially. As can be seen in Table 4 (M4), after doing so, the ΔCFI ($= −0.009$) indicated partial scalar invariance for the two-factor model. The intercept for item 2 of the positive affect scale (“in good spirits”) was higher in the USA ($b = 3.481$) than in Iran ($b = 3.282$). For item 6 (“full of life”), the intercept was lower in the USA ($b = 2.870$) than in Iran ($b = 3.208$).

**Latent Mean Differences**

Given that at least two indicators (including the referent indicators) for each of the factors were scalar invariant (Byrne et al., 1989), we proceeded with comparing latent means across the groups, using the parameters of the last model (M4). In this model, the latent factor means in Iran were constrained to zero, whereas the latent means in the USA were freely estimated. The results showed that the USA scored significantly higher than Iran on the frequency of positive affect (unstandardized fitted mean$_{USA} = 0.237$, $SE = 0.019$, $p < .001$). Additionally, the USA scored lower

---

**Table 1. Fit indexes for single-group confirmatory factor analyses**

<table>
<thead>
<tr>
<th>Culture</th>
<th>$\chi^2$</th>
<th>$df$</th>
<th>CFI</th>
<th>SRMR</th>
<th>RMSEA</th>
<th>90% CI for RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iran</td>
<td>726.295</td>
<td>51</td>
<td>0.945</td>
<td>0.043</td>
<td>0.074</td>
<td>0.070 – 0.079</td>
</tr>
<tr>
<td>USA</td>
<td>624.355</td>
<td>51</td>
<td>0.962</td>
<td>0.038</td>
<td>0.072</td>
<td>0.067 – 0.077</td>
</tr>
</tbody>
</table>

© 2015 Hogrefe Publishing

than Iran on the frequency of negative affect (unstandardized fitted mean (USA) = 1.069, \( \text{SE} = 0.026, \ p < .001 \)).

### Analyses Across Gender Groups

#### Single-Group CFAs

The two-correlated-factor model with residual covariances provided an acceptable fit in females (\( \chi^2 = 732.474, \ df = 51, \ p < .001 \). CFI = 0.956, RMSEA = 0.076) and males (\( \chi^2 = 762.679, \ df = 51, \ p < .001 \). CFI = 0.951, RMSEA = 0.079) in the whole sample, including Iranian and US participants. This model was used as the baseline model in the subsequent analyses.

#### Measurement Invariance

The baseline model with no equality constraint was simultaneously tested across gender groups. As can be seen in Table 3 (M5), the fit of this model to the data was...
acceptable, indicating configural invariance. Equality constraints were then imposed on all factor loadings. As shown in Table 3 (M6), the ΔCFI (= −0.001) indicated full metric invariance. Finally, equality constraints were imposed on all item intercepts to test scalar invariance (M7). The ΔCFI (= −0.002) indicated full scalar invariance.

### Latent Mean Differences

In the last model (M7), the latent factor means in females were constrained to zero, whereas the latent means in males were freely estimated. Using the whole data consisting of both national samples, gender groups did not differ significantly on the frequency of positive (unstandardized fitted mean(male) = 0.035, SE = 0.019, p = .06) and negative affect (unstandardized fitted mean(male) = −0.008, SE = 0.028, p = .782). Two separate latent mean analyses in each country showed that, in Iran, the two genders did not differ significantly on the frequency of positive (unstandardized fitted mean(male) = 0.044, SE = 0.028, p = .11) and negative affect (unstandardized fitted mean(male) = −0.009, SE = 0.037, p = .810). In the USA, the two genders did not differ significantly on the frequency of positive affect (unstandardized fitted mean(male) = 0.043, SE = 0.024, p = .078). However, the American men scored significantly lower than the American women on negative affect (unstandardized fitted mean(male) = −0.111, SE = 0.025, p < .001).

### Discussion

#### Factor Structure

The results indicate that the two-correlated-factor model was consistent with the data across culture and gender groups. The study also found that the two-factor model fitted the data better than a single-factor model. These findings are in line with the findings of prior research showing that positive and negative affect are distinct yet correlated components rather than two ends of a single continuum (Schimmack, 2008).

Specifying a covariance between the residuals of items 2 and 3 of the negative affect scale largely improved the fit of the model. This residual covariance seems justifiable given the content of the items. Items 2 (“nervous”) and 3 (“restless and fidgety”) both seem to be related to anxiety, whereas the other four items of this scale are more closely related to depression (e.g., hopeless, sad, worthless). For the positive affect scale, specifying a covariance between the residual terms of items 1 and 2 largely improved the fit of the model. Items 1 (“cheerful”), 2 (“in good spirits”), and 3 (“extremely happy”) all seem to capture a happy state of mind, whereas the other three items of this scale capture physical or more passive positive emotions (i.e., calm and peaceful, satisfied, full of life). Items 1 and 2 seem to be distinguishable from item 3 in that although all the three items assess an active and happy state of mind, items 1 and 2 seem to capture a state of moderate intensity compared to item 3, which indicates high intensity given the word “extremely.” Therefore, items 1 and 2 both capture active but not necessarily high intensity positive emotionality.

#### Cross-Cultural Results

Given full metric and partial scalar invariance, it can be concluded that the constructs of positive and negative affect as measured by these scales are largely equivalent across Iran and the USA. This indicates that, the 10 invariant items are interpreted and responded to similarly in Iran and the USA. This indicates that the 10 invariant items are interpreted and responded to similarly in Iran and the USA. It is important to note that, full measurement invariance seems scientifically unrealistic and, thus, partial invariance is a more realistic goal in practical research (e.g., Steinmetz, 2011). Therefore, that 10 out of 12 items of the scales are invariant across these two remarkably different cultures is promising.

This study revealed that the intercept for item 2 of the positive affect scale (“in good spirits”) was higher in the USA than in Iran. For item 6 (“full of life”), the intercept was lower in the USA than in Iran. Speculating about why the intercept for item 2 was not invariant between the two groups is difficult at this point given the dearth of studies in Iran, and remains a fruitful avenue for future studies. Regarding item 6, it can be speculated that the reason

---

**Table 3. Testing for measurement invariance**

<table>
<thead>
<tr>
<th>Across countries</th>
<th>(\chi^2)</th>
<th>df</th>
<th>CFI</th>
<th>RMSEA</th>
<th>(M) comparison</th>
<th>ΔCFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1. Configural invariance</td>
<td>1,350.650</td>
<td>102</td>
<td>0.955</td>
<td>0.073</td>
<td>– –</td>
<td>– –</td>
</tr>
<tr>
<td>M2. Full metric invariance</td>
<td>1,575.466</td>
<td>112</td>
<td>0.947</td>
<td>0.076</td>
<td>M2-M1</td>
<td>−0.008</td>
</tr>
<tr>
<td>M3. Full scalar invariance</td>
<td>2,249.945</td>
<td>122</td>
<td>0.923</td>
<td>0.088</td>
<td>M3-M2</td>
<td>−0.024</td>
</tr>
<tr>
<td>M4. Partial scalar invariance</td>
<td>1,824.033</td>
<td>120</td>
<td>0.938</td>
<td>0.079</td>
<td>M4-M2</td>
<td>−0.009</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Across gender groups</th>
<th>(\chi^2)</th>
<th>df</th>
<th>CFI</th>
<th>RMSEA</th>
<th>(M) comparison</th>
<th>ΔCFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>M5. Configural invariance</td>
<td>1,495.153</td>
<td>102</td>
<td>0.954</td>
<td>0.078</td>
<td>– –</td>
<td>– –</td>
</tr>
<tr>
<td>M6. Full metric invariance</td>
<td>1,515.823</td>
<td>112</td>
<td>0.953</td>
<td>0.074</td>
<td>M2-M1</td>
<td>−0.001</td>
</tr>
<tr>
<td>M7. Full scalar invariance</td>
<td>1,593.278</td>
<td>122</td>
<td>0.951</td>
<td>0.073</td>
<td>M3-M2</td>
<td>−0.002</td>
</tr>
</tbody>
</table>

**Note.** \(M\) = model.
why the intercept for item 6 was lower in the USA is that this item is related to the feeling of physical energy and activity, the frequency of which is expected to decrease with age. Considering that the American sample in this study is much older than the Iranian sample, this intercept difference is to be expected.

Although the measurement model was largely invariant across the samples, there were significant mean differences. The results of latent mean analysis revealed that the Iranian sample scored lower than the American sample on the frequency of positive affect, and significantly higher on the frequency of negative affect. These findings are consistent with the predictions of the present study that the more collectivistic and less wealthy Iran would score lower on positive affect and higher on negative affect. Prior cross-cultural research shows that social contexts in wealthy Western cultures promote more positive emotions, while in non-Western cultures, other affective states (e.g., wariness and shame) are also emphasized and promoted (Eid & Diener, 2001; Mesquita & Albert, 2007). For example, in the USA, “failure to achieve happiness ... can be seen as one of the greatest failures a person can experience” (Morris, 2012, p. 436). In contrast, in Iranian-Islamic culture, expressions of happiness and being very happy are sometimes condemned (Joshanloo, 2013a; Joshanloo & Weijers, 2014), and negative emotions are sometimes praised and promoted (Good & Good, 1988). Emotional states that conflict with cultural norms may be deemphasized and, accordingly, experienced less frequently (Diener, Oishi, & Ryan, 2013). Given different cultural and religious norms and lower individualism and socioeconomic development in Iran, it is not surprising that positive emotions are less frequently and negative emotions are more frequently experienced in this country.

However, it should be noted that age differences between the two cultural groups might have affected these results. Previous research has shown that the frequency of positive affect generally increases with age, and the frequency of negative affect generally decreases with age (Mroczek & Kolarz, 1998). Therefore, it is reasonable to suggest that the differences in latent means might have been partially caused by the fact that the US participants were older than the Iranian participants.

Cross-Gender Results

The scales possessed full metric and full scalar invariance across gender groups, which indicates that the construct and items are understood and responded to largely similarly by women and men across both nations. When both samples were lumped together, no latent mean differences were found between gender groups. However, the results of latent mean analyses in individual cultures indicated that there was a significant gender difference in the American sample, such that the American men scored significantly lower than the American women on the frequency of negative affect. This finding is in line with the prediction of the present study and previous findings in Western nations, showing that females generally score higher than males on the frequency of negative emotions (Nolen-Hoeksema & Rusting, 2003). This gender difference may be partly explained by such factors as hormonal differences between the genders, women's more vulnerability to social and relationship stress, and women's higher emphatic involvement with other's problems (Nolen-Hoeksema & Rusting, 2003).

However, this gender difference was not observed in the Iranian sample, a finding that is not fully explainable at the moment given the scarcity of relevant research in Iran. Nonetheless, this finding is consistent with some prior cross-cultural findings. For example, Fischer and Manstead's (2000) review of available cross-cultural studies suggests that “gender differences in emotions are more pronounced in Western Countries than in non-Western countries” (p. 74). It is likely that American culture emphasizes more strongly emotional expressiveness for women and emotional stability for men than Iranian culture, a speculation that merits further investigation.

Final Remarks

The present study has some strengths, such as its large samples from two nations. However, there are several limitations that should be addressed in future research. For example, although the Iranian sample is large, it cannot be seen as nationally representative, because it largely represents Iranian young adults. Moreover, there are remarkable differences between the two national samples used in this study. Besides the considerable age difference, different sampling methods have been used in the MIDUS and IMWP. Whereas random sampling has been used for data collection in the USA, the Iranian sample has been collected using convenience sampling. Given that this study was the first to investigate the factor structure of the Mroczek and Kolarz's scales of positive and negative affect, and in light of these limitations, we encourage future efforts to replicate these results using representative samples from more nations.

In sum, this study unraveled striking similarities in the structure of positive and negative affect in two markedly different cultures, and across gender groups. The results support full invariance across gender groups, and indicate that the majority of items of the affect scales function similarly across the cultural groups. These results confirm the validity of Mroczek and Kolarz’s scales for assessing positive and negative affect in these two cultures. Besides providing fresh evidence on cross-cultural and cross-gender measurement invariance of these scales, the present study also contributes to the nascent field of research on mental well-being in Iran. Consistent with several past studies using translations of Western well-being scales in Iran (as reviewed in Joshanloo, 2013b), this study indicates that Western well-being scales can be expected to work reasonably well within the Iranian context.
References


Walen, H. R., & Lachman, M. E. (2000). Social support and strain from partner, family, and friends: Costs and benefits